

Métodos basados en el supuesto de Independencia Condicional

Econometría IAE

Ricardo Pasquini

Noviembre 2023

Sesgo en diferencia *naïve*

- ▶ Anteriormente, analizamos una comparación naive para estimar el resultado causal de una intervención (T) sobre un outcome (Y).
- ▶ Obtuimos que la diferencia ex-post entre afectados por una intervención y los no afectados es igual al Tratamiento sobre los Tratados más un sesgo de selección:

$$E[Y_i \mid T_i = 1] - E[Y_i \mid T_i = 0] = \text{ATT} + \text{Preexistent bias}$$

Sesgo en diferencia *naïve*

- ▶ El sesgo lo pensamos como una diferencia de los resultados contrafácticos (de no recibir la intervención) entre tratados y no tratados:

$$E[Y_{0i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 0]$$

- ▶ Previamente vimos que podemos eliminar este sesgo si realizamos una asignación aleatoria del tratamiento. ¿Qué hacemos si no tenemos esta opción?

CIA - Supuesto de Independencia Condicional

- ▶ El Supuesto de Independencia Condicional (CIA) afirma que, condicional a una serie de características observadas X_i , los resultados potenciales son independientes del tratamiento

$$Y_{0i}, Y_{1i} \perp\!\!\!\perp T_i \mid X_i$$

Por lo tanto, condicional a X tenemos que el sesgo de selección desaparece:

$$E[Y_{0i} \mid X_i, T_i = 1] = E[Y_{0i} \mid X_i, T_i = 0]$$

CIA - Supuesto de Independencia Condicional

- ▶ En otras palabras, para unidades que suponemos que son comparables (ya que presentan un mismo valor de X) si diferenciamos entre participantes y no participantes de un programa, dará como resultado el efecto del tratamiento:

$$E[Y_i | X_i, T_i = 1] - E[Y_i | X_i, T_i = 0] = E[Y_{1i} - Y_{0i} | X_i, T_i = 1]$$

- ▶ Importante: Estamos suponiendo que todo lo que condiciona la toma del tratamiento se captura es capturable por los observables X !

Matching

- ▶ ¿Cómo podemos recuperar el ATT si se cumple CIA?
- ▶ Si podemos estimar un efecto para cada valor de X entonces podemos recuperar haciendo un promedio ponderado de los efectos.
- ▶ Para estimar el efecto para un valor de X solo necesitamos tratados y no tratados para ese valor de X (matching para X)

$$\delta_{ATT} = E[Y_{1i} - Y_{0i} \mid T_i = 1]$$

$$= E[E[Y_{1,i} - Y_{0,i} \mid X_i, T_i = 1], T_i = 1]$$

X CIA

$$E[Y_{0i} \mid X_i, T_i = 0] = E[Y_{0i} \mid X_i, T_i = 1]$$

Matching

Por lo tanto

$$\delta_{ATT} = E[E[Y_i | X_i, T_i = 1] - E[Y_i | X_i, T_i = 0] | T_i = 1]$$

$$= E[\delta_X | T_i = 1]$$

donde δ_X es el efecto de específico para un valor de X .

$$\delta_X = E[Y_i | X_i, T_i = 1] - E[Y_i | X_i, T_i = 0]$$

Matching. El estimador

- ▶ En el caso que X toma valores discretos vale que

$$ATT = E[Y_{1i} \mid T_i = 1] - E[Y_{0i} \mid T_i = 1] = \sum_x \delta_x P(X_i = x \mid T_i = 1)$$

- ▶ Por lo tanto, un estimador para el ATT es

$$\hat{ATT} = \sum_x \hat{\delta}_x P(\hat{X}_i = \hat{x} \mid T_i = 1)$$

CASO: Dale y Krueger 2002

ESTIMATING THE PAYOFF TO ATTENDING A MORE SELECTIVE COLLEGE: AN APPLICATION OF SELECTION ON OBSERVABLES AND UNOBSERVABLES*

STACY BERG DALE AND ALAN B. KRUEGER

Estimates of the effect of college selectivity on earnings may be biased because elite colleges admit students, in part, based on characteristics that are related to future earnings. We matched students who applied to, and were accepted by, similar colleges to try to eliminate this bias. Using the College and Beyond data set and National Longitudinal Survey of the High School Class of 1972, we find that students who attended more selective colleges earned about the same as students of seemingly comparable ability who attended less selective schools. Children from low-income families, however, earned more if they attended selective colleges.

Figure 1: Dale y Krueger

CASO: Dale y Krueger 2002

"Past studies have found that students who attended colleges with higher average SAT scores or higher tuition tend to have higher earnings when they are observed in the labor market. Attending a college with a 100 point higher average SAT is associated with 3 to 7 percent higher earnings later in life (see, e.g., Kane [1998]). As Kane notes, an obvious concern with this conclusion is that students who attend more elite colleges may have greater earnings capacity regardless of where they attend school."

CASO: Dale y Krueger 2002 (simplificado)

TABLE 2.1
The college matching matrix

Applicant group	Student	Private			Public			1996 earnings
		Ivy	Leafy	Smart	All State	Tall State	Altered State	
A	1		Reject	Admit		Admit		110,000
	2		Reject	Admit		Admit		100,000
	3		Reject	Admit		Admit		110,000
B	4	Admit			Admit		Admit	60,000
	5	Admit			Admit		Admit	30,000
C	6		Admit					115,000
	7		Admit					75,000
D	8	Reject			Admit	Admit		90,000
	9	Reject			Admit	Admit		60,000

Note: Enrollment decisions are highlighted in gray.

Figure 2: College Matrix

CASO: Dale y Krueger 2002 (simplificado)

Ver que:

- ▶ La diferencia de privada vs publica me da aprox 20,000 (92 vs 72,5 en promedio)
- ▶ La diferencia si matcheamos por grupo y ponderamos es:

$$\frac{3}{5}(-5.000) + \frac{2}{5}(30.000) = 9.000$$

- ▶ Notar que al matchear tengo que descartar los grupos C y D

CASO: Dale y Krueger 2002 (simplificado)

TABLE I
ILLUSTRATION OF HOW MATCHED-APPLICANT GROUPS WERE CONSTRUCTED

Student	Matched-applicant group	Student applications to college							
		Application 1		Application 2		Application 3		Application 4	
		School average	School admissions	School average	School admissions	School average	School admissions	School average	School admissions
SAT	decision	SAT	decision	SAT	decision	SAT	decision	SAT	decision
Student A	1	1280	Reject	1226	Accept*	1215	Accept	na	na
Student B	1	1280	Reject	1226	Accept	1215	Accept*	na	na
Student C	2	1360	Accept	1310	Reject	1270	Accept*	1155	Accept
Student D	2	1355	Accept	1316	Reject	1270	Accept*	1160	Accept
Student E	2	1370	Accept*	1316	Reject	1260	Accept	1150	Accept
Student F	Excluded	1180	Accept*	na	na	na	na	na	na
Student G	Excluded	1180	Accept*	na	na	na	na	na	na
Student H	3	1360	Accept	1308	Accept*	1260	Accept	1160	Accept
Student I	3	1370	Accept*	1311	Accept	1255	Accept	1155	Accept
Student J	3	1350	Accept	1316	Accept*	1265	Accept	1155	Accept
Student K	4	1245	Reject	1217	Reject	1180	Accept*	na	na
Student L	4	1235	Reject	1209	Reject	1180	Accept*	na	na
Student M	5	1140	Accept	1055	Accept*	na	na	na	na
Student N	5	1145	Accept*	1060	Accept	na	na	na	na
Student O	No match	1370	Reject	1038	Accept*	na	na	na	na

* Denotes school attended.

na = did not report submitting application.

The data shown on this table represent hypothetical students. Students F and G would be excluded from the matched-applicant subsample because they applied to only one school (the school they attended). Student O would be excluded because no other student applied to an equivalent set of institutions.



Digresión: CIA y CEF

- ▶ Notar que una manera alternativa de estimar

$$E[Y_i | X_i, T_i = 1] - E[Y_i | X_i, T_i = 0]$$

es aplicando el modelo de regresión!!

Podemos implementar un CEF perfectamente cuando las categorías de la variable X son finitas.

- ▶ En este caso, el supuesto sería

$$E[Y_i | X_i, T_i] = \beta_0 + \beta_1 X + \beta_2 T + \epsilon$$

- ▶ Si no tenemos categorías finitas, igualmente debiéramos encontrar resultados similares. La diferencia entre los dos métodos es que la regresión pondera las observaciones de acuerdo a la varianza relativa, mientras el matching de acuerdo al número de observaciones (Véase MHE.)

CASO: Dale y Krueger 2002 (simplificado)

- ▶ En el caso de Dale y Krueger, el modelo de regresión implementado es:

$$\log(\text{earnings}) = \alpha_i + \beta_1 \text{Private}_i + \beta_2 SAT + \epsilon_i$$

CASO: Dale y Krueger 2002 (simplificado)

TABLE 2.2
Private school effects: Barron's matches

From Mastering Metrics: The Path from Cause to Effect © 2013 Princeton University Press. Used by permission.
All rights reserved.

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	.135 (.055)	.095 (.052)	.086 (.034)	.007 (.038)	.003 (.039)	.013 (.025)
Own SAT score ÷ 100		.048 (.009)	.016 (.007)		.033 (.007)	.001 (.007)
Log parental income			.219 (.022)			.190 (.023)
Female				−.403 (.018)		−.395 (.021)
Black				.005 (.041)		−.040 (.042)
Hispanic				.062 (.072)		.032 (.070)
Asian				.170 (.074)		.145 (.068)
Other/missing race				−.074 (.157)		−.079 (.156)
High school top 10%				.095 (.027)		.082 (.028)
High school rank missing				.019 (.033)		.015 (.037)
Athlete				.123 (.025)		.115 (.027)
Selectivity-group dummies	No	No	No	Yes	Yes	Yes

Notes: This table reports estimates of the effect of attending a private college or university on earnings. Each column reports coefficients from a regression of log earnings on a dummy for attending a private institution and controls. The results in columns (4)–(6) are

Propensity Score Matching (PSM)

- ▶ Propensity Score Matching (PSM) construye un grupo de comparación estadística que se basa en un modelo de probabilidad de participar en el tratamiento, utilizando características observadas.
- ▶ La idea es emparejar a cada individuo del grupo de tratamiento con un individuo del grupo de control con una probabilidad similar de participar en el programa.

Propensity Score Matching (PSM)

Esto implica dos supuestos:

- ▶ CIA: Todo lo que puede explicar la condición de tratamiento se debe a factores observables.
 - ▶ En esto PSM es similar a la regresión (controlando por observables) y al matching.
- ▶ Presencia de un soporte común o superposición, que garantice que podré encontrar un individuo “similar” en el grupo de control. Esto es:

$$0 < P(T_i = 1|X_i) < 1$$

Notar que si

$$P(T_i = 1|X_i) = 1$$

, todos tomarían el programa.

Propensity Score Matching (PSM) - Soporte común

Figure 4.1 Example of Common Support

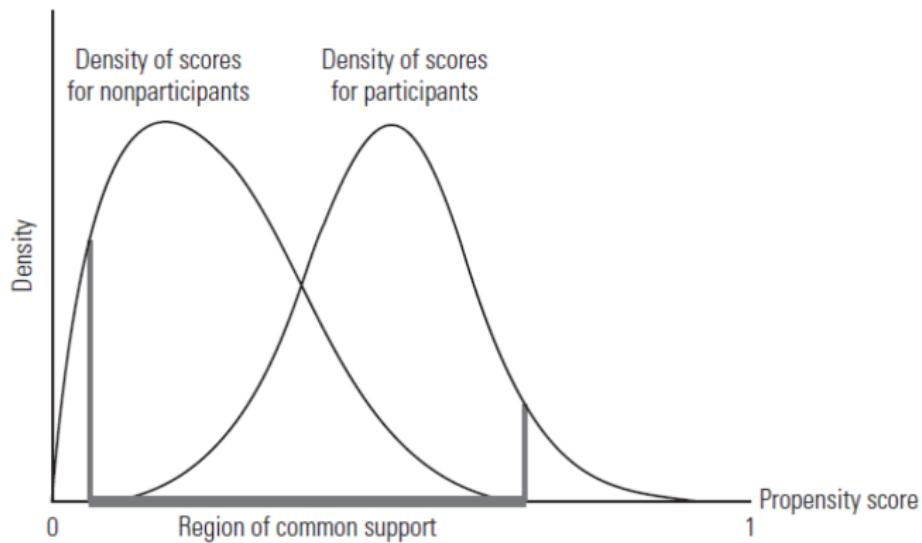


Figure 5: 1565205953462

Propensity Score Matching (PSM)

En la práctica, la estimación se realiza en tres pasos:

- ▶ Primero, se estima $p(T_i = 1|X_i)$ utilizando algún tipo de modelo paramétrico, por ejemplo, Logit o Probit.
- ▶ Segundo, se emparejan individuos tratados y no tratados sobre la base de las probabilidades estimadas $\hat{p}(T_i = 1|X_i)$ (Ver métodos más adelante)
- ▶ Tercero, las estimaciones del ATT se calculan evaluando las diferencias.

$$\text{ATT}_{PSM} = \frac{1}{N_T} \left[\sum_{i \in T} Y_i^T - \sum_{i \in T} w(i, j) Y_i^C \right]$$

donde $w(i, j)$ son pesos cuando hay más de un control por cada observación tratada.

Propensity Score Matching (PSM)

- ▶ Matching entre tratados y no tratados:
 - ▶ Vecino más cercano (nearest neighbor matching): simplemente tomar los n vecinos más cercanos sobre la base de $p(T_i = 1|X_i)$.
 - ▶ Calibre o radio (Caliper or radius matching): es posible que no haya vecinos de larga distancia, luego imponga un radio o umbral.
 - ▶ Estratificación o emparejamiento de intervalos. Divide el soporte común en estratos (o intervalos) y calcula el impacto del programa dentro de cada intervalo.
 - ▶ Kernel Matching: tiene en cuenta todos los matchs posibles y los pondera según la distancia.

Propensity Score Matching (PSM)

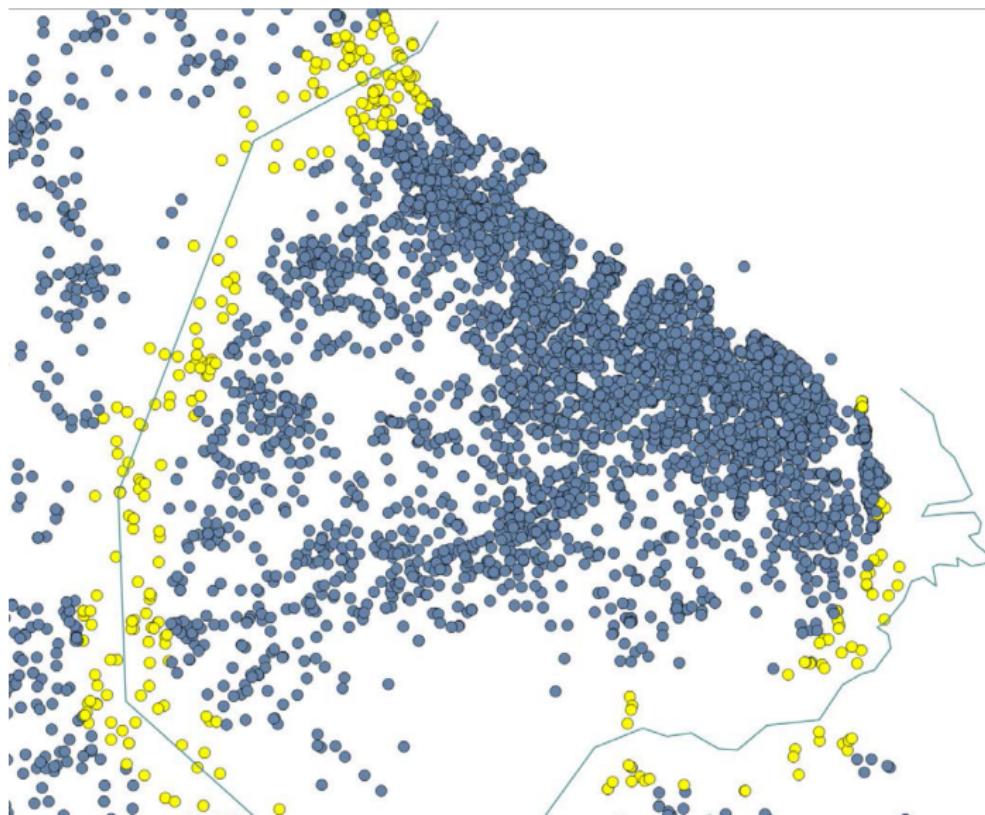
- ▶ Cálculo de errores estándar mediante Bootstrap
- ▶ Usando propensity scores como pesos en un enfoque de regresión. Hirano, Imbens y Ridder (2003) Estimación:
$$Y_i = \alpha + \beta T_i + \rho X_i + \varepsilon_i$$
 con pesos de 1 para los participantes y pesos de $\hat{P}(X)/(1 - \hat{P}(X))$ para las observaciones de control.

Caso: Efectos ley de alquileres

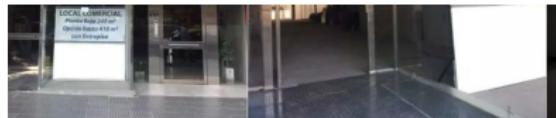
- ▶ El 8 de Septiembre de 2017 la legislatura de la Ciudad de Buenos Aires dictaminó la ley 5859, que entró en vigencia el 16 de septiembre. Esta ley fue sancionada con el objetivo de beneficiar a los inquilinos. La ley estableció, en primer lugar, que la comisión inmobiliaria no sería pagada por los inquilinos, sino por los dueños. Esto modificó lo que era la práctica habitual, en la cual se cargaba al inquilino con el costo. Esto implicaría una reducción en el costo de entrada a un contrato de alquiler para los inquilinos, ya que estos en muchos casos solían pagar un mes, o más en términos de comisión. Además, la ley estableció un máximo para esta comisión, correspondiente a un mes extra (el equivalente a 4,15% -1 mes sobre 24- del valor total del contrato)
- ▶ Idea de la investigación: Identificar si hubo un incremento en los alquileres causado por la ley.
- ▶ Una versión relacionada en Pasquini (2020, JHE)

Caso: Efectos ley de alquileres

- ▶ Estrategia de Identificación: Enfoque cuasi-experimental.
 - ▶ Matching con datos post-intervención.



Caso: Efectos ley de alquileres



Ejemplo Matching results 1

Proprietà - Capital Federal - Nuñez - Local - Alquiler

ALQUILER

Local comercial en Venta / Alquiler - Av. Cabildo 4700 - Nuñez - Local Comercial En Venta / Alquiler - Av. Cabildo 4700 - Nuñez - \$ 115.000
Locales en Alquiler en Nuñez, Capital Federal, Argentina



Proprietà - Baja G.B.A. Zona Norte - Nueve Locas - Florida - Local - Alquiler

ALQUILER

Único Local Comercial, Apto Todo Destino sobre Av. Maipú - Florida - Maipú Al 800 - \$ 150.000
Locales en Alquiler en Florida, Vicente López, Bs.As. G.B.A. Zona Norte, Argentina

Detalle Mapa StreetView Fotos en galería

► Descripción

PLANTA BAJA/200m2 cubiertos en planta baja totalmente libres. Pisos acceso vehicular con puerta especial concreta (en 3 planos) para facilitar la carga y descarga de mercaderías, así como, el acceso de automóviles, motos y camionetas al interior del salón ya que cuenta con una altura de techo superior a 3mts. Esta

Contactar al anfitrión: Carlos Castillo Propiedades
Nueve

Figure 7: 1565207339294

Caso: Efectos ley de alquileres

Ejemplo Matching results 2

Properati · Bs.As. G.B.A. Zona Norte · Vicente López · Florida · Departamento · Alquiler ·

Depto 2 Amb Exc Vista a 2/C de Maipu - Florida - Melo Al 1600 - \$ 7.000
Departamentos en Alquiler en Florida, Vicente López, Bs.As. G.B.A. Zona Norte, Argentina

Este aviso ha caducado

Properati · Capital Federal · Nuñez · Departamento · Alquiler ·

Unico dos ambientes en Torre Astor con cochera y baulera. - Nuñez - Pico Al 2300 Astor Núñez - \$ 16.500
Departamentos en Alquiler en Nuñez, Capital Federal, Argentina

Este aviso ha caducado

Figure 8: 1565207368637



Ejemplo Matching results 3